Statistical Analysis of Solar Neutrino Data

M.V. Garzelli and **C. Giunti**

INFN, Sezione di Torino
and
Dipartimento di Fisica Teorica,
Università di Torino
Via P. Giuria 1, I-10125 Torino, Italy

garzelli@to.infn.it
giunti@to.infn.it

Standard Method

Least-Squares estimator of Δm^2 and θ: $X^2_{\min}$

$$X^2 = \sum_{j_1,j_2} \left(R^{(\mathrm{thr})}_{j_1} - R^{(\mathrm{exp})}_{j_1} \right) (V^{-1})_{j_1 j_2} \left(R^{(\mathrm{thr})}_{j_2} - R^{(\mathrm{exp})}_{j_2} \right)$$

$R^{(\mathrm{thr})}_j$ = theoretical rate for experiment or bin j
$R^{(\mathrm{exp})}_j$ = rate measured in experiment or bin j
$j = 1, \ldots, N_{\mathrm{exp}}$, where N_{exp} is the number of data points
V = covariance matrix: experimental and theoretical uncertainties

Standard Goodness of Fit

Probability to observe a minimum of X^2 larger than the one actually observed, assuming for $X^2_{\min}$ a χ^2 distribution with $N_{\mathrm{dof}} = N_{\mathrm{exp}} - N_{\mathrm{par}}$ degrees of freedom (N_{par} is the number of fitted parameters).

Standard Allowed Regions

The standard $100\beta\%$ CL regions in the $\tan^2\vartheta$–Δm^2 plane are given by the condition

$$X^2 \leq X^2_{\min} + \Delta X^2(\beta)$$

β = Confidence Level (CL)
$\Delta X^2(\beta)$ = value of χ^2 such that the cumulative χ^2 distribution for 2 degrees of freedom is equal to β

$$\beta = 90\% \ (1.64\,\sigma) \Rightarrow \Delta X^2(0.90) = 4.61$$
$$\beta = 99\% \ (2.58\,\sigma) \Rightarrow \Delta X^2(0.99) = 9.21$$

Analysis of Rates

Rates of Homestake [1], GALLEX+SAGE [2, 3], Super-Kamiokande 2001 [4] $\Rightarrow N_{\text{exp}} = 3$, $N_{\text{par}} = 2$, $N_{\text{dof}} = 1$.

Global Analysis

Rates of Homestake [1], GALLEX+SAGE [2, 3], Super-Kamiokande 2001 [4]: 3 data points

$$+$$

Shape of Super-Kamiokande 2000 [5] Day-Night data: 6 bins and 1 normalization factor

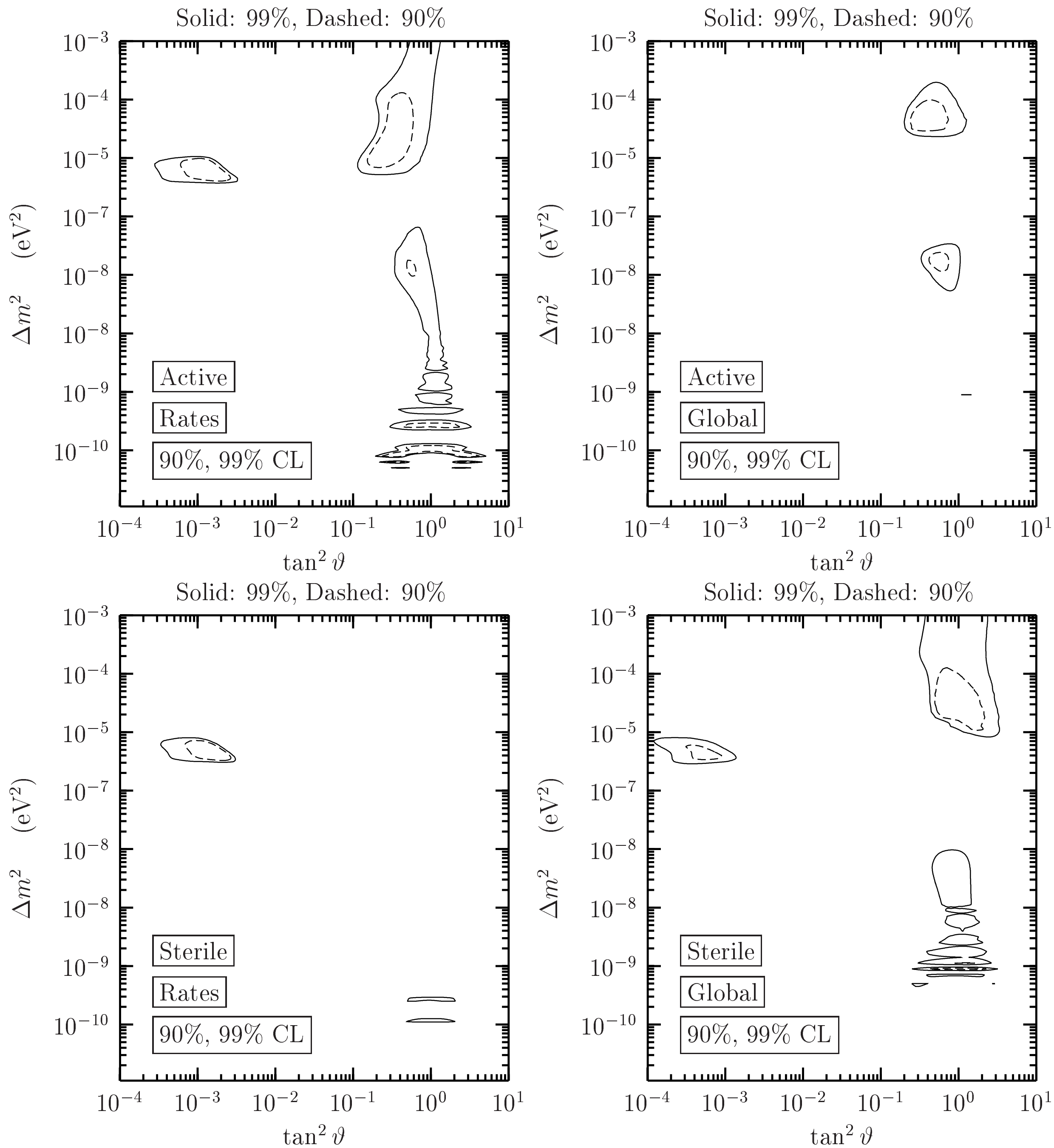
$$+$$

Shape of Super-Kamiokande 2001 [4] energy spectrum for $E > 5.5\,\text{MeV}$: 18 bins and 1 normalization factor

$$\Downarrow$$

$$N_{\text{exp}} = 27,\ N_{\text{par}} = 4,\ N_{\text{dof}} = 23.$$

$$\boxed{\begin{array}{l}\text{Active} \Rightarrow \nu_e \to \nu_{\mu,\tau} \\ \text{Sterile} \Rightarrow \nu_e \to \nu_s\end{array}}$$



Best fit values in Table at pag. 11.

NOVELTIES

Super-Kamiokande rate/BP-SSM has decreased from 0.474 ± 0.020 [6] to 0.451 ± 0.008 [4].

Active Rates

Best Fit in VO region! Due to decrease of SK rate/BP-SSM.

Active Global

Best Fit continues to be in LMA region!
No SMA region at 99% CL! Due to flat spectrum.
At 99% CL VO region almost vanish! Due to flat energy spectrum and Day-Night asymmetry (albeit small).

Sterile Rates

Best Fit continues to be in SMA region!

Sterile Global

Poor Goodness of Fit $\Longrightarrow$ sterile disfavored!
Best Fit in LMA region! Due to flat energy spectrum and decrease of Super-Kamiokande rate/BP-SSM (the incompatibility of a flat energy spectrum with the Homestake rate is alleviated).

Conditions for the validity of the Standard Method

$\boxed{\mathbf{1}}$ The theoretical rates $R_j^{(\mathrm{thr})}$ depend $\underline{\text{linearly}}$ on the parameters Δm^2 and $\tan^2\theta$ to be determined in the fit.

$\boxed{\mathbf{2}}$ The errors $R_j^{(\mathrm{thr})} - R_j^{(\mathrm{exp})}$ are $\underline{\text{multinormally}}$ distributed.

$\boxed{\mathbf{3}}$ The covariance matrix V does not depend on Δm^2 and $\tan^2\theta$.

In reality these three conditions are $\underline{\text{not}}$ satisfied:

$\boxed{\mathbf{1}}$ The theoretical rates $R_j^{(\mathrm{thr})}$ do not depend at all linearly on the parameters Δm^2, $\tan^2\theta$. This is the reason why there are several allowed regions in the $\tan^2\theta$–Δm^2 plane and these regions do not have elliptic form.

$\boxed{\mathbf{2}}$ The errors $R_j^{(\mathrm{thr})} - R_j^{(\mathrm{exp})}$ are not multinormally distributed, because although the fluxes ϕ_i^{SSM} and the cross sections $C_{ij}^{(\mathrm{thr})}$ are assumed to be multinormally distributed, their products, that determine the theoretical rates through the relations

$$R_j^{(\mathrm{thr})} = \sum_i C_{ij}^{(\mathrm{thr})} \, \phi_i^{\mathrm{SSM}} \,, \tag{1}$$

are not multinormally distributed.

$\boxed{\mathbf{3}}$ The covariance matrix V depends on Δm^2 and $\tan^2\theta$.

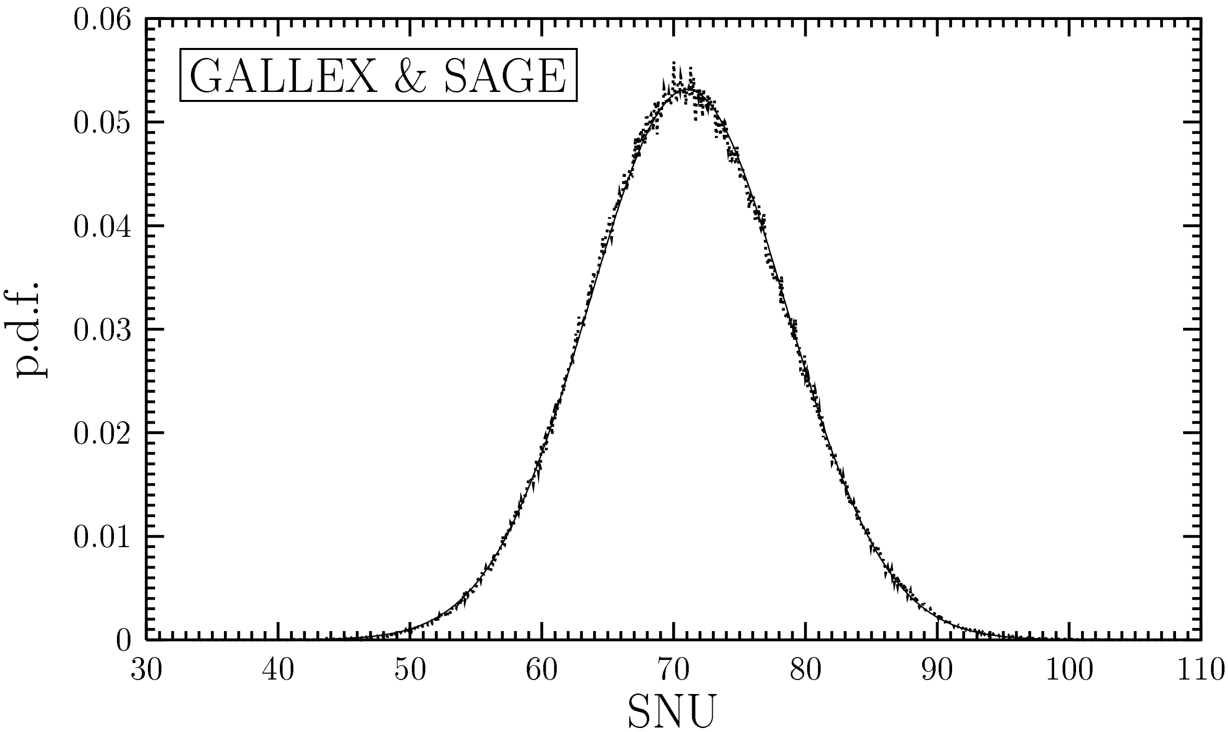
$\boxed{\mathbf{1}}$ is <u>important</u>!

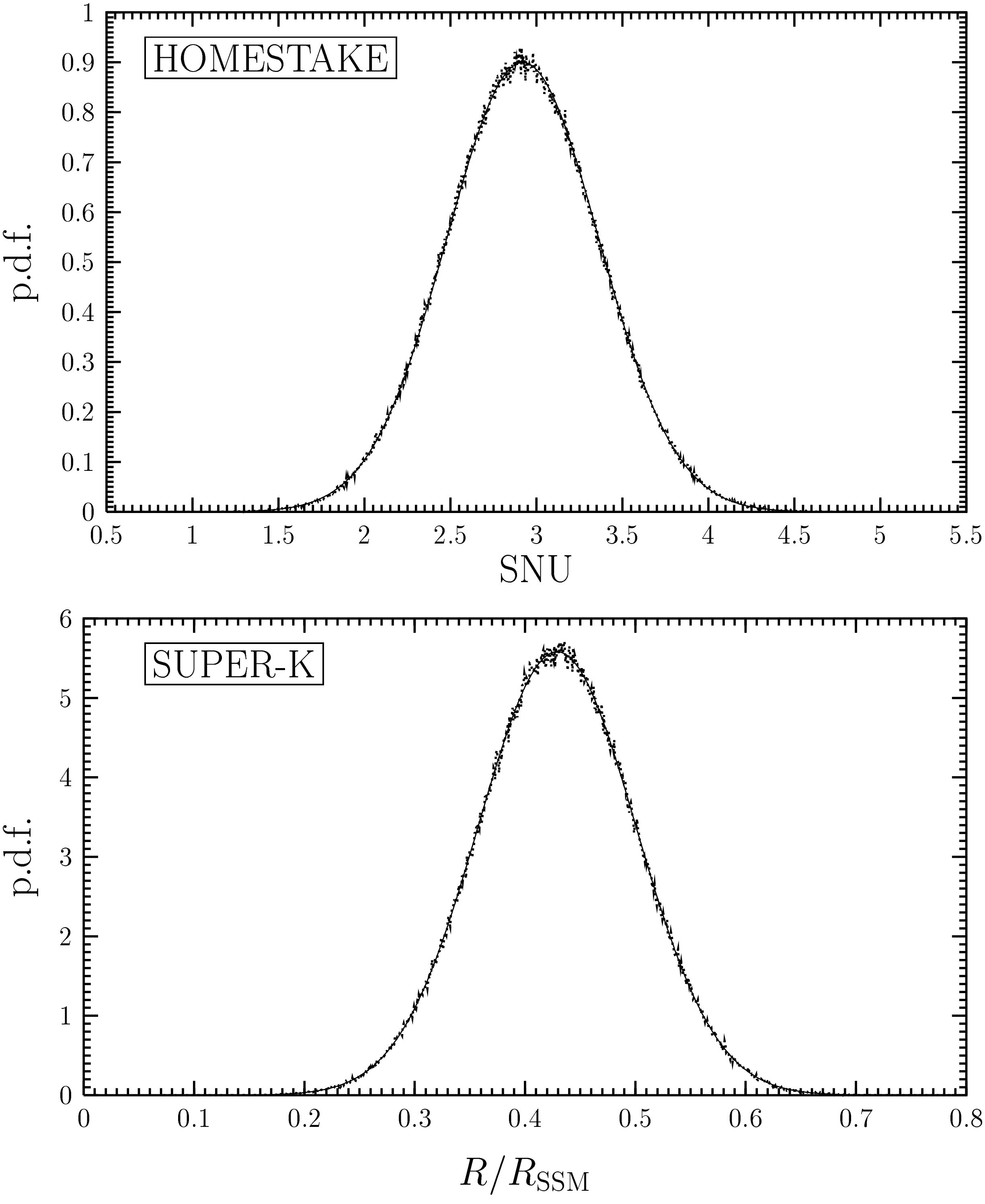
$\boxed{\mathbf{2}}$ is <u>irrelevant</u>: the multinormal approximation is very good, as shown by the following three figures.

Dotted lines: probability distribution function of experimental rates generated with Monte Carlo.

Solid lines: probability distribution function of experimental rates assuming a multinormal distribution given by the Likelihood function

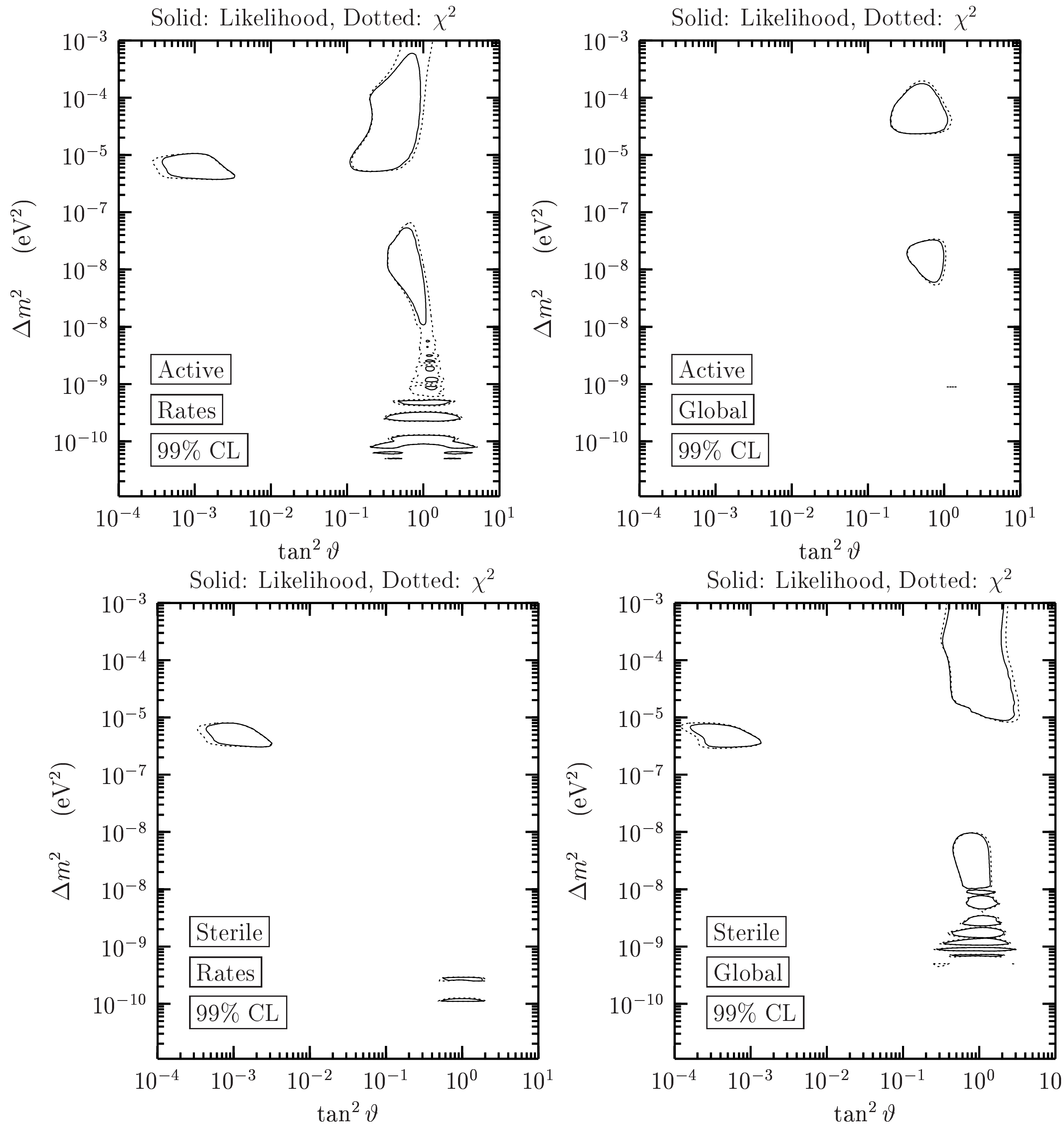
$$\mathcal{L}(R_j^{(\exp)} | \tan^2\theta, \Delta m^2) = \frac{e^{-X^2/2}}{(2\pi)^{N_{\exp}/2}\sqrt{|V|}}$$




HOMESTAKE
p.d.f.
SNU
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
SUPER-K
p.d.f.
R/R_{SSM}
0
1
2
3
4
5
6
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8

$\boxed{\mathbf{3}}$ is not negligible, as shown by the following comparison of the Standard Regions with the regions obtained with the log-Likelihood method (see [7])

$$\ln \mathcal{L} \geq \ln \mathcal{L}_{\max} - \frac{\Delta X^2(\beta)}{2}$$



Monte Carlo Goodness of Fit [8]

▶ Estimate best-fit values of Δm^2, $\tan^2\theta$ through $X^2_{\rm min}$.

▶ Call the best-fit values $\widehat{\Delta m^2}$, $\widehat{\tan^2\theta}$.

▶ Assume that $\widehat{\Delta m^2}$, $\widehat{\tan^2\theta}$ are reasonable surrogates of the true values $\Delta m^2_{\rm true}$, $\tan^2\theta_{\rm true}$.

▶ Using $\widehat{\Delta m^2}$, $\widehat{\tan^2\theta}$, generate N_s synthetic random data sets with the standard gaussian distribution for the experimental and theoretical uncertainties.

▶ Apply the Least-Squares method to each synthetic data set, leading to an ensemble of simulated best-fit parameters $\widehat{\Delta m^2}_{(s)}$, $\widehat{\tan^2\theta}_{(s)}$ with $s = 1, \ldots, N_s$, each one with his associated $(X^2_{\rm min})_s$.

▶ Calculate GoF as the fraction of simulated $(X^2_{\rm min})_s$ in the ensemble that are larger than the one actually observed, $X^2_{\rm min}$.

		Goodness of Fit	
		Standard	Monte Carlo
Active Rates	$X^2_{\text{min}} \simeq 1.2$ $\widehat{\Delta m^2} \simeq 9 \times 10^{-11}\,\text{eV}^2$ $\widehat{\tan^2\vartheta} \simeq 0.45$	27%	8%
Active Global	$X^2_{\text{min}} \simeq 24$ $\widehat{\Delta m^2} \simeq 5 \times 10^{-5}\,\text{eV}^2$ $\widehat{\tan^2\vartheta} \simeq 0.40$	40%	37%
Sterile Rates	$X^2_{\text{min}} \simeq 0.97$ $\widehat{\Delta m^2} \simeq 4 \times 10^{-6}\,\text{eV}^2$ $\widehat{\tan^2\vartheta} \simeq 2 \times 10^{-3}$	32%	21%
Sterile Global	$X^2_{\text{min}} \simeq 37$ $\widehat{\Delta m^2} \simeq 3 \times 10^{-5}\,\text{eV}^2$ $\widehat{\tan^2\vartheta} \simeq 0.8$	3%	2%

$\rightsquigarrow$ Standard method overestimates GoF!

$\rightsquigarrow$ **Explanation**: GoF is the probability to obtain better fits than the observed one.
When there are several local minima of X^2 with relatively close values of X^2, there are more possibilities to obtain good fits and the true goodness of fit is smaller than the standard one (obtained assuming that X^2 is a quadratic function of $\tan^2\vartheta$, Δm^2, with only one minimum).

Monte Carlo CL of Standard Allowed Regions [8]

Definition: $100\beta\%$ CL Allowed Regions belong to a set of allowed regions that cover the true value of the parameters with probability β.

$\lightning$ Given the usual "$100\beta\%$ CL" allowed regions in the $\tan^2\vartheta$–Δm^2 plane, calculate their Monte Carlo Confidence Level β_{MC} with a method similar to the one used for the Goodness of Fit.

► Assume that $\widehat{\Delta m^2}$, $\widehat{\tan^2\theta}$ are reasonable surrogates of the true values Δm^2_{true}, $\tan^2\theta_{\text{true}}$.

► Generate a large number of synthetic data sets.

► Apply the standard procedure to each synthetic data set and obtain the corresponding "$100\beta\%$ CL" Standard Allowed Regions in the $\tan^2\theta$–Δm^2 plane.

► Count the number of synthetic "$100\beta\%$ CL" Standard Allowed Regions that cover the assumed surrogate $\widehat{\Delta m^2}$, $\widehat{\tan^2\theta}$ of the true values.

► The ratio of this number and the total number of synthetically generated data set gives the Confidence Level β_{MC} of the "$100\beta\%$ CL" Standard Allowed Regions.

	CL of Standard Allowed Regions	
	Standard	Monte Carlo
Active Rates	90.00% ($1.64\,\sigma$)	84.08% ($1.41\,\sigma$)
	99.00% ($2.58\,\sigma$)	98.16% ($2.36\,\sigma$)
Active Global	90.00% ($1.64\,\sigma$)	82.83% ($1.37\,\sigma$)
	99.00% ($2.58\,\sigma$)	97.56% ($2.25\,\sigma$)
Sterile Rates	90.00% ($1.64\,\sigma$)	87.11% ($1.52\,\sigma$)
	99.00% ($2.58\,\sigma$)	98.51% ($2.43\,\sigma$)
Sterile Global	90.00% ($1.64\,\sigma$)	80.01% ($1.28\,\sigma$)
	99.00% ($2.58\,\sigma$)	97.09% ($2.18\,\sigma$)

$\rightsquigarrow$ The Confidence Level of Standard Allowed Regions is smaller than its nominal value.

$\rightsquigarrow$ **Explanation**: When there are several local minima of X^2 with relatively close values of X^2, in repeated experiments the global minimum has significant chances to occur far from the true value of $\tan^2\vartheta$ and Δm^2 $\Rightarrow$ smaller probability that the allowed regions cover the true value of $\tan^2\vartheta$ and Δm^2, with respect to the case in which there is only one minimum.

Frequentist Allowed Regions

$\Rightarrow$ **Frequentist Statistics** allows to calculate allowed regions with correct **coverage** using Neyman's method.

$\nRightarrow$ But there is arbitrariness in the choice of

1) Estimator of the parameters

2) Method for the construction of acceptance regions

↯ In [8] we have calculated "exact" confidence regions using as estimate of $\tan^2\vartheta$ and Δm^2 their value at X^2_{min}.

⅄ In [9] it has been argued that X^2_{min} may be an insufficient estimator, leading to a loss of information. Notice that if this is true, the standard χ^2 method suffers from the same problem!

⅄ In order to prevent any loss of information, it is better to use the full data set as estimator of $\tan^2\vartheta$ and Δm^2, as done in [9].

$\nRightarrow$ However, there is still the problem of choice of the method for the construction of acceptance intervals.

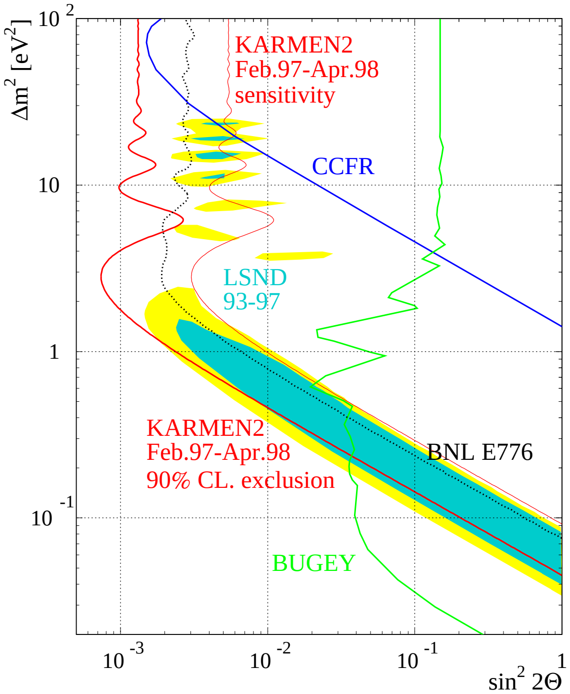
⅄ In [9] it has been argued that the Unified Approach (UA) [10] is more appropriate than the smallest acceptance intervals method, also known as "Crow–Gardner" (CG).

$\nRightarrow$ Unfortunately, it is well known that when the UA differs from the smallest acceptance intervals method it gives unreliable confidence intervals (see [11, 12])

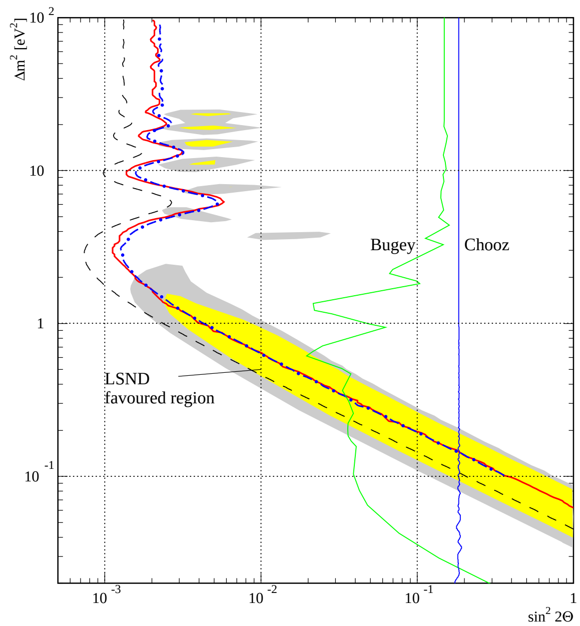
$\nRightarrow$ Infamous example:

The **KARMEN** 1998 limit on $\bar{\nu}_\mu \to \bar{\nu}_e$ oscillations obtained with the Unified Approach was unreliably much more stringent than the sensitivity of the experiment [13].

The **KARMEN** 1999 limit is less stringent than the 1998 one. More data $\Longrightarrow$ less information!



KARMEN 1998 exclusion limit and sensitivity [13].



KARMEN 1999 exclusion limit and sensitivity [14]. Solid line: 1999 limit. Dash-Dotted line: 1999 sensitivity. Dashed line: 1998 limit.

$\not\Rightarrow$ Other infamous examples:

$\not\Rightarrow$ The 1999 limit on neutrinoless double-beta decay obtained in the **Heidelberg-Moscow** experiment [15] obtained with the Unified Approach was much more stringent than the sensitivity of the experiment. That is why now they do not use the Unified Approach any more [16]! (Got burned!)

$\not\Rightarrow$ The present upper limit on $\nu_\mu \to \nu_\tau$ neutrino oscillations obtained in the NOMAD experiment [17] is stronger than the one obtained in the CHORUS experiment [18] not because the NOMAD experiment has a better sensitivity than the CHORUS experiment (see discussion in [18]), but because the NOMAD collaboration uses the Unified Approach, which gives unphysically stringent upper bounds when the number of observed events is smaller than the expected background.

↯ The UA and similar methods [19, 20, 21] are appropriate in order to get allowed regions even in the presence of an unlikely statistical fluctuation of the data, such that the data are very unlikely for any value of the parameters.

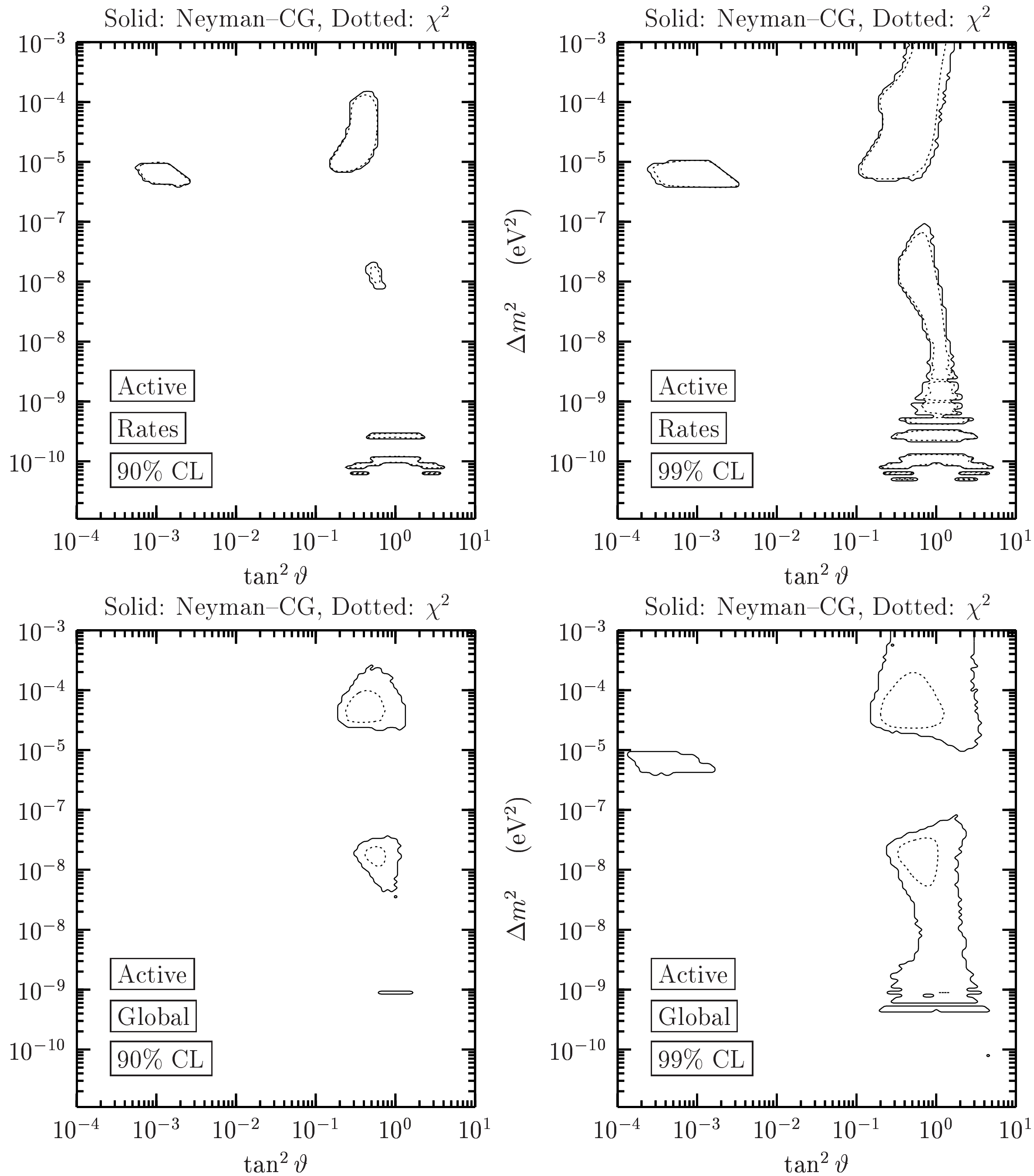
However, the physical reliability of such allowed regions is highly questionable.

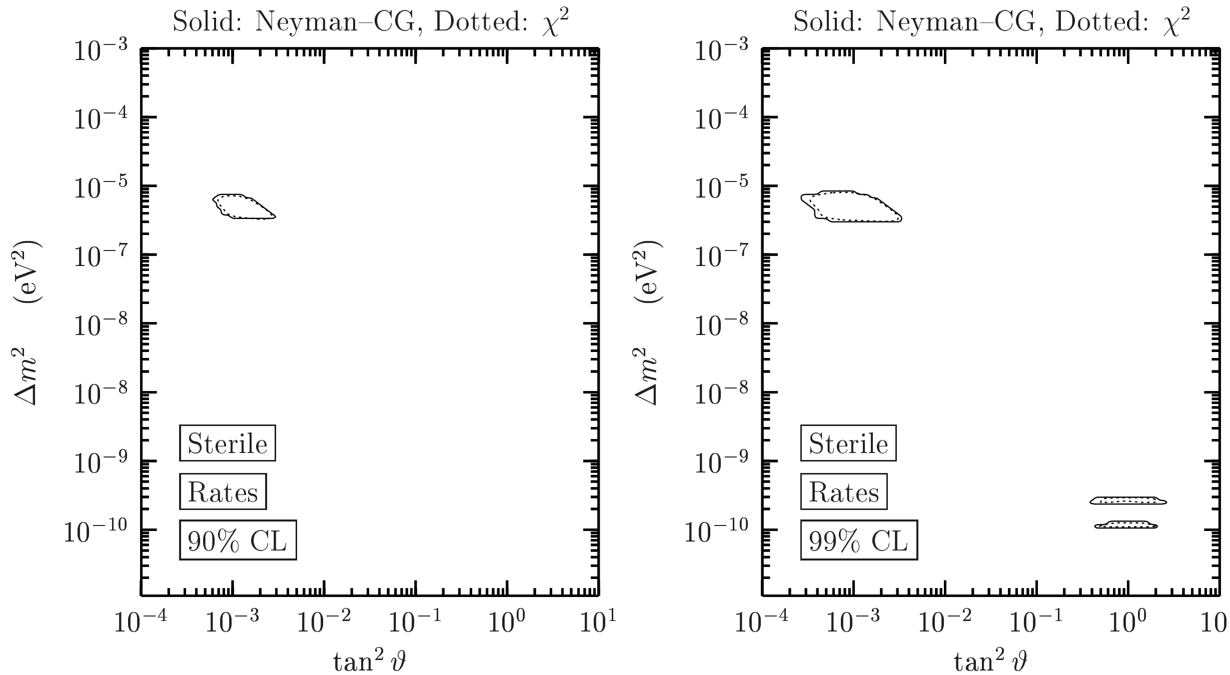
↯ If there is no statistical fluctuation of data, the UA and the CG methods are equivalent.

↯ From the value of the GoF (see Table at pag. 11) one can see that **there is no unlikely statistical fluctuation in solar neutrino data** in the case of $\nu_e \to \nu_{\mu,\tau}$ oscillations and in the case of the analysis of the rates in terms of $\nu_e \to \nu_s$ oscillations.

On the other hand, if the solar neutrino problem is due to $\nu_e \to \nu_s$ oscillations, there is an unlikely statistical fluctuation of the shape of the energy spectrum and the global analysis of solar ν data with the CG method is unreliable.

↯ Therefore, the CG method (that is computationally much easier than the UA method) can be applied to the analisis of solar ν data in terms of $\nu_e \to \nu_{\mu,\tau}$ oscillations and to the analysis of the rates of solar neutrino experiments in terms of $\nu_e \to \nu_s$ oscillations.


Solid: Neyman–CG, Dotted: χ^2
Δm^2 (eV^2)
$\tan^2\vartheta$
10^{-3}
10^{-4}
10^{-5}
10^{-7}
10^{-8}
10^{-8}
10^{-9}
10^{-10}
10^{-4} 10^{-3} 10^{-2} 10^{-1} 10^{0} 10^{1}
Active
Rates
90% CL
Active
Rates
99% CL
Active
Global
90% CL
Active
Global
99% CL



Solid: Neyman–CG, Dotted: χ^2
Sterile
Rates
90% CL
Solid: Neyman–CG, Dotted: χ^2
Sterile
Rates
99% CL
Δm^2 (eV^2)
$\tan^2\vartheta$

Bayesian Allowed Regions

⇒ Bayesian Theory allows to calculate the **improvement of knowledge as a consequence of experimental measurements** (see [22]).

⇒ This is how our mind works and how science improves. Therefore, **Bayesian Theory is the natural statistical tool for scientists**.

▶ Bayesian probability density function of $\tan^2\vartheta$ and Δm^2 after measurement of rates $R_j^{(\mathrm{exp})}$:

$$p(\tan^2\theta, \Delta m^2 | R_j^{(\mathrm{exp})}) \propto \mathcal{L}(R_j^{(\mathrm{exp})} | \tan^2\theta, \Delta m^2) p(\tan^2\theta, \Delta m^2)$$

$p(\tan^2\theta, \Delta m^2)$ = prior probability density function

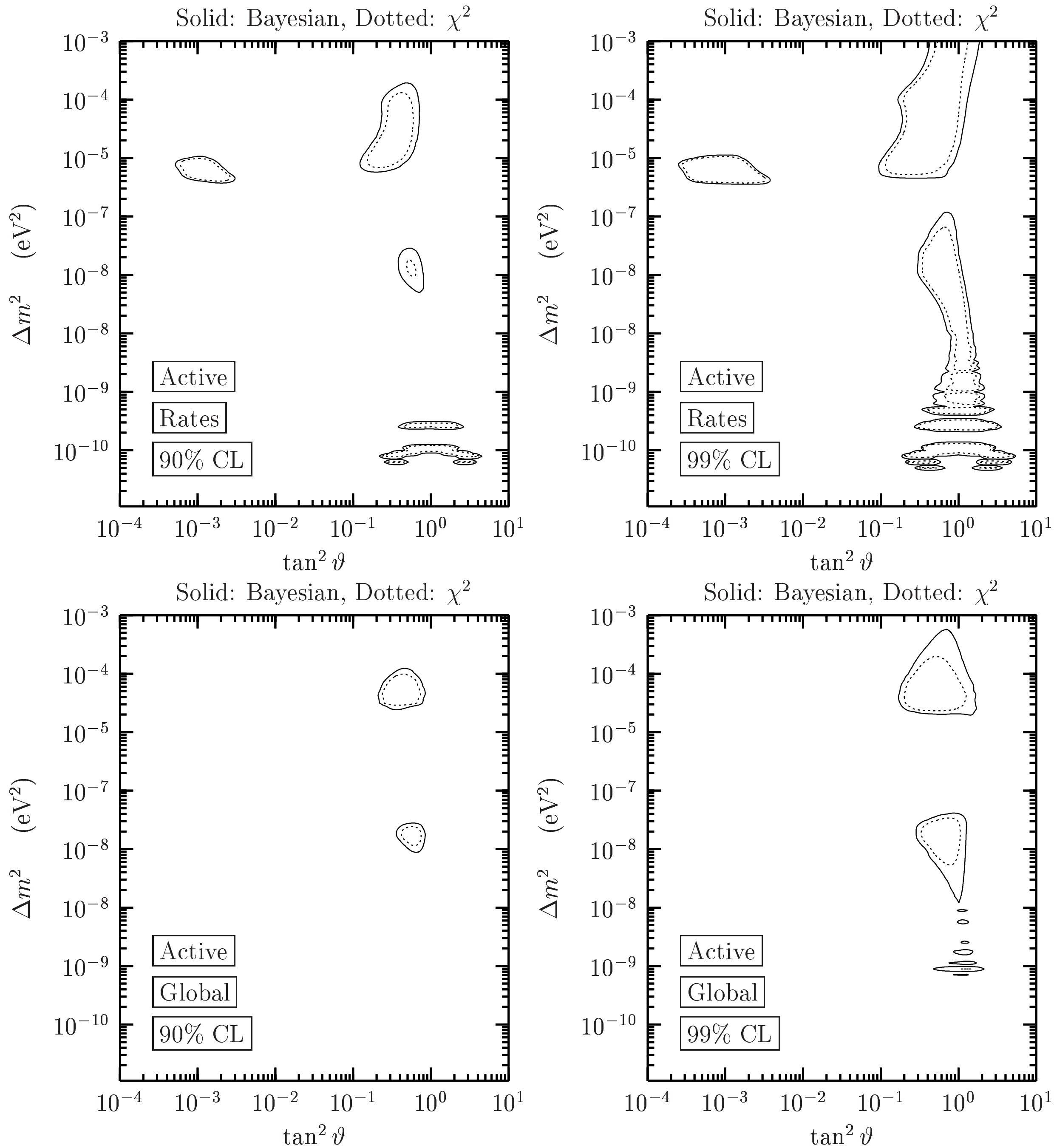
↯ Prior knowledge on $\tan^2\vartheta$ and Δm^2: All values are allowed, but we know that solar ν data are sensitive to different orders of magnitude of $\tan^2\vartheta$ and Δm^2 through different mechanisms.

$$\Downarrow$$

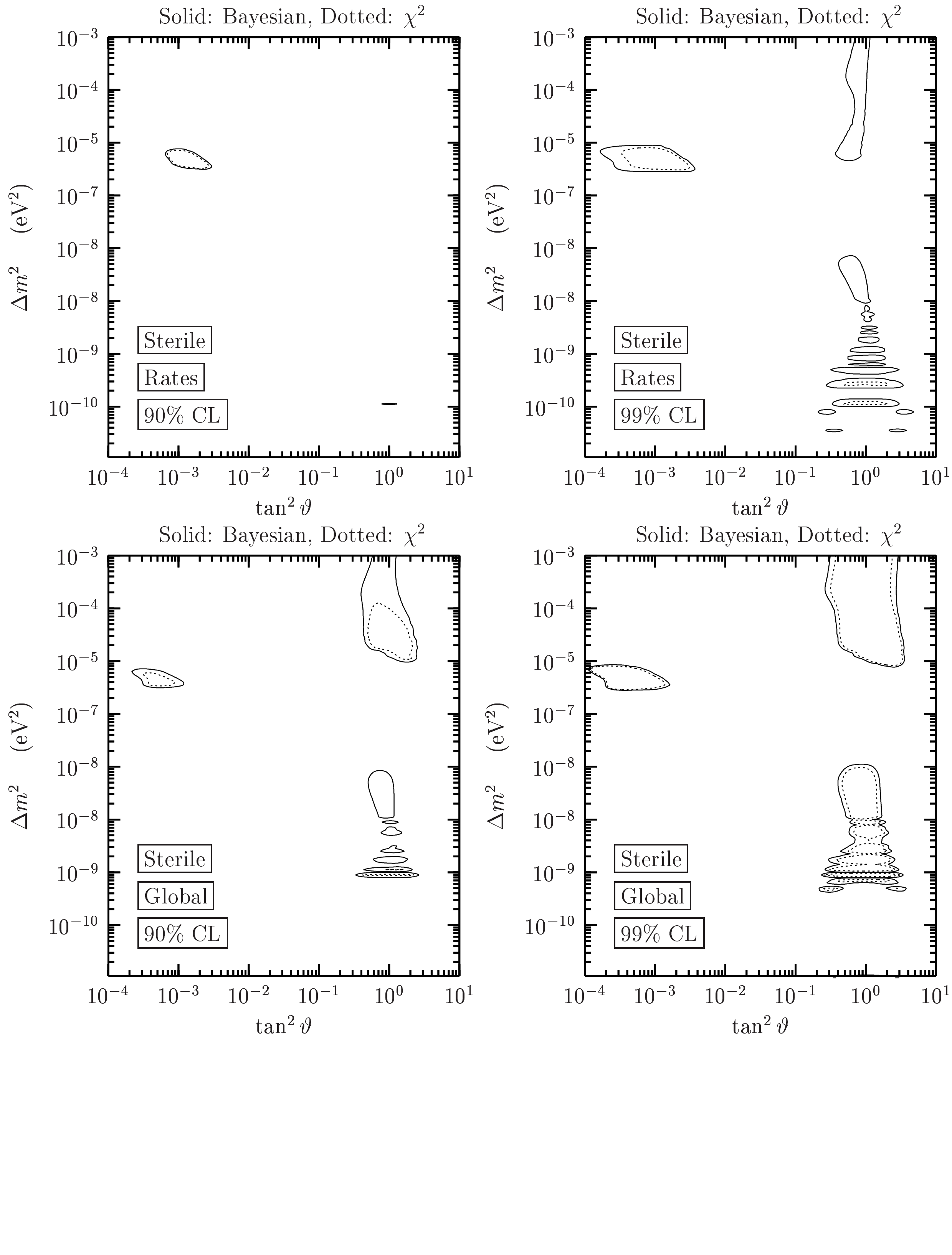
Flat prior in the $\log(\tan^2\vartheta)$–$\log(\Delta m^2)$ plane

$$\Downarrow$$

$$\boxed{\begin{aligned} & p(\tan^2\theta, \Delta m^2 | R_j^{(\mathrm{exp})}) = \\ & = \frac{\mathcal{L}(R_j^{(\mathrm{exp})} | \tan^2\theta, \Delta m^2)}{\int \mathcal{L}(R_j^{(\mathrm{exp})} | \tan^2\theta, \Delta m^2) \, \mathrm{d}\log(\tan^2\vartheta) \, \mathrm{d}\log(\Delta m^2)} \end{aligned}}$$



Solid: Bayesian, Dotted: χ^2
Δm^2 (eV^2)
$\tan^2\vartheta$
10^{-4} 10^{-3} 10^{-2} 10^{-1} 10^{0} 10^{1}
10^{-3} 10^{-4} 10^{-5} 10^{-7} 10^{-8} 10^{-8} 10^{-9} 10^{-10}
Active
Rates
90% CL
99% CL
Global


Solid: Bayesian, Dotted: χ^2
Sterile
Rates
90% CL
Sterile
Rates
99% CL
Sterile
Global
90% CL
Sterile
Global
99% CL
Δm^2 (eV^2)
$\tan^2\vartheta$

References

[1] B. T. Cleveland *et. al.*, "Measurement of the solar electron neutrino flux with the homestake chlorine detector," *Astrophys. J.* **496** (1998) 505.

[2] **GALLEX** Collaboration, W. Hampel *et. al.*, "Gallex solar neutrino observations: Results for gallex iv," *Phys. Lett.* **B447** (1999) 127.

[3] **SAGE** Collaboration, J. N. Abdurashitov *et. al.*, "Measurement of the solar neutrino capture rate with gallium metal," *Phys. Rev.* **C60** (1999) 055801, `astro-ph/9907113`.

[4] **SuperKamiokande** Collaboration, S. Fukuda *et. al.*, "Solar b-8 and hep neutrino measurements from 1258 days of super-kamiokande data," `hep-ex/0103032`.

[5] Y. Suzuki (Super-Kamiokande Coll.), Talk presented at Neutrino 2000, Sudbury, Canada, 16–21 June 2000 (http://nu2000.sno.laurentian.ca); Talk presented at NOW2000, Otranto, Italy, September 2000 (http://www.ba.infn.it/˜now2000).

[6] **Super-Kamiokande** Collaboration, Y. Fukuda *et. al.*, "Measurement of the solar neutrino energy spectrum using neutrino electron scattering," *Phys. Rev. Lett.* **82** (1999) 2430–2434, `hep-ex/9812011`.

[7] W.T. Eadie, D. Drijard, F.E. James, M. Roos and B. Sadoulet, *Statistical Methods in Experimental Physics*, North Holland, Amsterdam, 1971.

[8] M. V. Garzelli and C. Giunti, "A frequentist analysis of solar neutrino data," `hep-ph/0007155`.

[9] P. Creminelli, G. Signorelli, and A. Strumia, "Frequentist analyses of solar neutrino data," `hep-ph/0102234`.

[10] G. J. Feldman and R. D. Cousins, "A unified approach to the classical statistical analysis of small signals," *Phys. Rev.* **D57** (1998) 3873–3889, `physics/9711021`.

[11] C. Giunti and M. Laveder, "The physical significance of confidence intervals," `hep-ex/0002020`.

[12] C. Giunti and M. Laveder, "The power of confidence intervals," `hep-ex/0011069`.

[13] **KARMEN** Collaboration, K. Eitel *et. al.*, "The search for neutrino oscillations $\bar{\nu}_\mu \to \bar{\nu}_e$ with karmen," *Nucl. Phys. Proc. Suppl.* **77** (1999) 212, `hep-ex/9809007`.

[14] T. E. Jannakos, "The search for neutrino oscillations $\bar{\nu}_\mu \to \bar{\nu}_e$ with karmen," `hep-ex/9908043`.

[15] L. Baudis *et. al.*, "Limits on the majorana neutrino mass in the 0.1-ev range," *Phys. Rev. Lett.* **83** (1999) 41–44, `hep-ex/9902014`.

[16] H. V. Klapdor-Kleingrothaus *et. al.*, "Latest results from the heidelberg-moscow double-beta-decay experiment," `hep-ph/0103062`.

[17] **NOMAD** Collaboration, P. Astier *et. al.*, "Updated results from the ν_τ appearance search in nomad," *Phys. Lett.* **B483** (2000) 387–404.

[18] **CHORUS** Collaboration, E. Eskut *et. al.*, "New results from a search for $\nu_\mu \rightarrow \nu_\tau$ and $\nu_e \rightarrow \nu_\tau$ oscillation," *Phys. Lett.* **B497** (2001) 8–22.

[19] C. Giunti, "New ordering principle for the classical statistical analysis of poisson processes with background," *Phys. Rev.* **D59** (1999) 053001, `hep-ph/9808240`.

[20] S. Ciampolillo, "Small signal with background: Objective confidence intervals and regions for physical parameters from the principle of maximum likelihood," *Nuovo Cim.* **A111** (1998) 1415–1430.

[21] M. Mandelkern and J. Schultz, "The statistical analysis of gaussian and poisson signals near physical boundaries," *J. Math. Phys.* **41** (2000) 5701–5709, `hep-ex/9910041`.

[22] G. D'Agostini, "Bayesian reasoning in high energy physics. Principles and applications," *CERN Yellow Report* **99-03** (1999).